## CLAIMS

What is claimed is:

1. A method for transmitting a mixed media data stream in packets, including audio and video objects, between a sender and a receiver through a connection over a network, the method comprising:

monitoring, at the receiver, transmission characteristics of the connection between the server and the receiver;

estimating available bandwidth at the sender based upon the transmission characteristics of the connection monitored at the receiver;

allocating a global buffer for the mixed media data stream to be transmitted from the sender to the receiver as a function of the estimated available bandwidth at the sender;

pre-encoding a portion of each Video Object Plane (VOP) in the global buffer with respect to a quantization parameter (QP) of the VOP;

encoding the VOP in the global buffer based on the QP;

updating a rate distortion model based upon the QP and packet loss rate;

performing a frame skipping function after the VOP encoding; and

transmitting from the sender to the receiver the encoded video object plane in the global buffer at a regulated sender transmission rate from the sender as a function of the estimated available bandwidth at the sender.

2. The method as defined in Claim 1, further comprising:

receiving the encoded video object plane at the receiver from the connection;

demultiplexing the encoded video object plane into coded video and audio streams;

inputting the coded video and audio streams, respectively, into video and audio decoders;

inputting the decoded video and audio streams to a media mixer; and

inputting the mixed video and audio streams output from the media mixer to an output device.

3. The method as defined in Claim 1, wherein pre-encoding a portion of each VOP with respect to the QP of the VOP further comprises adjusting the QP of the VOP.

4. The method as defined in Claim 3, wherein the QP of the VOP is adjusted with respect to a texture parameter, r, as the number of bits which will be used to encode the VOP wherein:

$$r = \frac{p_1 \times MAD}{QP} + \frac{p_2 \times MAD}{QP^2};$$

$p_1$ and $p_2$ are control parameters; and

MAD is a mean absolute distortion in which a total target bit rate for all objects in the global buffer are allocated proportionally to motion, size, and square of MAD.

5. The method as defined in Claim 4, wherein the adjusting of the QP of the VOP is performed by changing the QP to a values in a range from 1 to 31 depending upon the estimated available bandwidth at the receiver.

6. The method as defined in Claim 1, wherein updating a rate distortion model based upon the QP and packet loss rate comprises:

predicting the number of bits, $r_i$, to encode the ith VOP, and is given by:

$$r_i = \frac{(p_1)_i \times MAD_i}{QP_i} + \frac{(p_2)_i \times MAD_i}{QP_i^2};$$

the distortion, d, is estimated by $d_i = (q_1)_i \times QP_i + (q_2)_i \times QP_i^2 + (q_3)_i \times r_i \times (P_L)_i$,

wherein:

$q_1$, $q_2$ and $q_3$ are control parameters; and

the packet loss rate $(P_L)_i$ is an estimate of that the probability that the ith transmission of data from the sender will be lost; and

minimizing the overall distortion, D, for each encoded VOP by $D = \sum_i d_i$, subject to $R = \sum_i r_i \le R_T$, where $R_T$ is the total bit budget for the current time instant.

7. The method as defined in Claim 1, wherein:

the sender sends data to the receiver in through a connection over a packet switched network in a sender packet having a sender header that includes:

a packet sequence number;

a timestamp indicating the time when the sender packet was sent (ST1); and

the size of the sender packet (PacketSize);

the receiver sends data to the sender through the connection over the packet switched network in a receiver packet having a receiver header that includes:

the time interval that the sender packet spent in the receiver side ($\Delta$RT);

the timestamp of the sender packet sent from the sender (ST1);

an estimate, calculated by the receiver, of a packet-loss rate; and

the rate at which data is received at the receiver;

monitoring transmission characteristics of the connection between server and receiver comprises:

estimating a round trip time of the sender packet from the sender to the receiver (RTT) based on ST1 and $\Delta$RT;

estimating a time out interval (TO) before which the sender should retransmit to the receiver a sender packet of data that has not been received by the receiver;

estimating a probability that a packet of data will be lost ($P_L$);

estimating the present available network bandwidth at which the receiver can receive data from the sender (rvcrate) as a function of the PacketSize, the RTT, the $P_L$, and the TO;

deriving the present sending rate of data from the sender to the receiver

($\overline{currate}$);

setting an updated sending rate of data from the sender to the receiver

(currate), wherein:

if rcvrate is greater than $\overline{currate}$, then deriving currate as a function $\overline{currate}$, PacketSize, and RTT; and

if rcvrate is not greater than $\overline{currate}$, then setting currate to be less than rcvrate.

8. The method as defined in Claim 7, wherein:

$RTT = \alpha \times \overline{RTT} + (1-\alpha) \times (now - ST1 - \Delta RT)$; and

$now$ is the timestamp indicating the time at which the receiver packet was received in the sender; and α is a weighting parameter.

9. The method as defined in Claim 7, wherein:

$TO = RTT + (k \times RTTVAR)$;

k is a constant;

$RTTVAR = \alpha_2 \times \overline{RTTVAR} + (1-\alpha_2) \times |RTT - (now - ST1 - \Delta RT)|$;

$\overline{RTTVAR}$ is the current variation in the round trip time of the sender packet from the sender to the receiver (RTT);

α2 is a weighting parameter; and

RTTVAR is a smoothed estimate of $\overline{RTTVAR}$.

10. The method as defined in Claim 7, wherein $P_L$ is derived by a Gilbert Model.

11. The method as defined in Claim 10, wherein:

$$P_L = \frac{\hat{q}}{\hat{p}+\hat{q}};$$

$$\{X_i\}_{i=1}^{n};$$

$X_i$ takes 1 if the ith sender packet has arrived successfully at the receiver;

$X_i$ takes 0 if the ith sender packet is lost;

$$p = P[X_i = 1 \mid X_{i-1} = 0];$$

$$q = P[X_i = 0 \mid X_{i-1} = 1];$$

$\hat{p}$ is an estimate of p;

$\hat{q}$ is an estimate of q; and

$\hat{p} = n_{01}/n_0$ and $\hat{q} = n_{10}/n_1$, wherein:

$n_{01}$ is the number of times in an observed time series when one follows zero;

$n_{10}$ is the number of times when zero follows one;

$n_0$ is the number of zeros; and

$n_1$ is the number of ones.

12. The method as defined in Claim 11, wherein:

the $P_L$ is further smoothed by a filter that weights the n most recent measured packet loss rates by:

$$P_{L,i} = \sum_{j=0}^{n-1}(w_j \times \overline{P_{L,i-j}})$$;

$\overline{P_{L,i-j}}$ is the measured packet loss rate in the (i-j)th time interval;

two set of weighting parameters are defined as follows:

| | W0 | W1 | W2 | W3 | W4 | W5 | W6 | W7 |
|---|---|---|---|---|---|---|---|---|
| WS1 | 1.0 | 1.0 | 1.0 | 1.0 | 0.8 | 0.6 | 0.4 | 0.2 |

| | W0 | W1 | W2 | W3 | W4 | W5 | W6 | W7 |
|---|---|---|---|---|---|---|---|---|
| WS2 | 1.2 | 1.2 | 1.0 | 1.0 | 0.8 | 0.5 | 0.3 | 0.1 |

; and WS2 is used for $w_j$ when the actual packet loss rate is less than half of the measured packet loss rate, otherwise WS1 is used for $w_j$.

13. A computer-readable media comprising computer-executable instructions for performing the method as recited in Claim 1.

14. A method for transmitting a mixed media data stream in packets, including audio and multiple video objects (MVOs), between a sender and a receiver through a connection over a network, the method comprising:

monitoring transmission characteristics of one or more encoded video object planes through the connection between the sender and the receiver;

estimating, from the transmission characteristics, an available bandwidth ($R_T$) at the sender;

allocating, as a function of the $R_T$, a portion of the mixed media data stream to a global buffer;

encoding a video object plane from the global buffer based upon a rate distortion function that accounts for packet loss rate between sender and receiver;

updating the rate distortion function based upon results of the encoded video object plane and upon a memory containing results of one or more previously encoded video object planes;

after the encoding the MVOs in the video object plane, performing a frame skipping function; and

transmitting, at the estimated available bandwidth, the encoded video object plane from the sender to the receiver.

15. The method as defined in Claim 14, wherein allocating a portion of the mixed media data stream to a global buffer comprises:

$W_{cur} = \max(((W_{prev} + B_{prev}) \times R_T / R_{old} - R_T / F),\ 0)$, as the global buffer size $R_{old}/2$, is changed to $R_T/2$, wherein:

$B_{prev}$ is the number of bits spent in the previous time instant $B_{prev}$,

$R_{old}/2$ is the previous size of the global buffer;

$W_{prev}$ is the previous occupancy of the global buffer; and

F is the video frame rate.

16. The method as defined in Claim 14, wherein allocating a portion of the mixed media data stream to a global buffer comprises the allocation of an output target rate from the global buffer among each of video and audio data streams so as to yield the target bits for an individual object in the data stream.

17. The method as defined in Claim 14, further comprising:

receiving the encoded video object plane at the receiver from the connection;

demultiplexing the encoded video object plane into coded video and audio streams;

inputting the coded video and audio streams, respectively, into video and audio decoders; and

inputting the decoded video and audio streams to a media mixer; and

inputting the mixed video and audio streams output from the media mixer to an output device.

18. A computer-readable media comprising computer-executable instructions for performing the method as recited in Claim 14.

19. A method for transmitting a mixed media data stream including multiple video objects (MVOs) between a sender and a receiver through a connection over a packet switched network, the method comprising:

transmitting from the sender a sender packet of data to the receiver, the sender packet including a sender header that includes:

a packet sequence number;

a timestamp indicating the time when the sender packet was sent (ST1); and

the size of the sender packet (PacketSize);

transmitting from the receiver a receiver packet of data to the sender, the receiver packet including a receiver header that includes:

the time interval that the sender packet spent in the receiver side ($\Delta$RT);

the timestamp of the sender packet sent from the sender (ST1);

an estimate, calculated by the receiver, of a packet-loss rate; and

the rate at which data is received at the receiver;

estimating a round trip time of the sender packet from the sender to the receiver (RTT) based on ST1 and $\Delta$RT;

estimating a time out interval (TO) before which the sender should retransmit to the receiver a sender packet of data that has not been received by the receiver;

estimating a probability that a packet of data will be lost ($P_L$);

estimating the present available network bandwidth at which the receiver can receive data from the sender (rcvrate) as a function of the PacketSize, the RTT, the $P_L$, and the TO;

deriving the present sending rate of data from the sender to the receiver ($\overline{currate}$);

setting an updated sending rate of data from the sender to the receiver (currate), wherein:

if rcvrate is greater than $\overline{currate}$, then deriving currate as a function $\overline{currate}$, PacketSize, and RTT; and

if rcvrate is not greater than $\overline{currate}$, then setting currate to be less than rcvrate.

20. The method as defined in Claim 19, wherein:

$RTT = \alpha \times \overline{RTT} + (1-\alpha) \times (now - ST1 - \Delta RT)$;

*now* is the timestamp indicating the time at which the receiver packet was received in the sender; and

α is a weighting parameter.

21. The method as defined in Claim 19, wherein:

$TO = RTT + (k \times RTTVAR)$;

k is a constant;

$RTTVAR = \alpha_2 \times \overline{RTTVAR} + (1-\alpha_2) \times |RTT - (now - ST1 - \Delta RT)|$;

$\overline{RTTVAR}$ is the current variation in the round trip time of the sender packet from the sender to the receiver (RTT);

α2 is a weighting parameter; and

RTTVAR is a smoothed estimate of $\overline{RTTVAR}$

22. The method as defined in Claim 19, wherein $P_L$ is derived by a Gilbert Model.

23. The method as defined in Claim 22, wherein:

$$P_L = \frac{\hat{q}}{\hat{p}+\hat{q}};$$

$$\{X_i\}_{i=1}^{n};$$

$X_i$ takes 1 if the ith sender packet has arrived successfully at the receiver;

$X_i$ takes 0 if the ith sender packet is lost;

$$p = P[X_t = 1 \mid X_{t-1} = 0];$$

$$q = P[X_t = 0 \mid X_{t-1} = 1];$$

$\hat{p}$ is an estimate of p;

$\hat{q}$ is an estimate of q; and

$\hat{p} = n_{01}/n_0$ and $\hat{q} = n_{10}/n_1$, wherein:

$n_{01}$ is the number of times in an observed time series when one follows zero;

$n_{10}$ is the number of times when zero follows one;

$n_0$ is the number of zeros; and

$n_1$ is the number of ones.

24. The method as defined in Claim 23, wherein:

the $P_L$ is further smoothed by a filter that weights the n most recent measured packet loss rates by $P_{L,i} = \sum_{j=0}^{n-1}(w_j \times \overline{P_{L,i-j}})$; $\overline{P_{L,i-j}}$ is the measured packet loss rate in the (i-j)th time interval;

two sets of weighting parameters are defined as follows:

| | W0 | W1 | W2 | W3 | W4 | W5 | W6 | W7 |
|---|---|---|---|---|---|---|---|---|
| WS1 | 1.0 | 1.0 | 1.0 | 1.0 | 0.8 | 0.6 | 0.4 | 0.2 |

| | W0 | W1 | W2 | W3 | W4 | W5 | W6 | W7 |
|---|---|---|---|---|---|---|---|---|
| WS2 | 1.2 | 1.2 | 1.0 | 1.0 | 0.8 | 0.5 | 0.3 | 0.1 |

; and WS2 is used for $w_j$ when the actual packet loss rate is less than half of the measured packet loss rate, otherwise WS1 is used for $w_j$.

25. The method as defined in Claim 24, wherein

$$rcvrate = \frac{PacketSize}{RTT \times \sqrt{2P_L/3} + 3 \times TO \times P_L \times \sqrt{3P_L/8} \times (1+32P_L{}^2)}.$$

26. The method as defined in Claim 19, wherein:

lastchange is the timestamp indicating the time at which last adjustment to currate occurred;

*now* is the present moment; and

setting currate further comprises:

if rcvrate is greater than $\overline{currate}$, then currate = $\overline{currate}$ + (PacketSize/RTT) X multi, where multi = (*now* - lastchange)/RTT, and multi is constrained from 1 to 2; and

if rcvrate is not greater than $\overline{currate}$, then currate = a constant X rcvrate + (1- the constant) X $\overline{currate}$.

27. The method as defined in Claim 19, further comprising:

allocating, as a function of the rvcrate, a portion of the mixed media data stream to a global buffer;

encoding a video object plane from the global buffer based upon a rate distortion function that accounts for packet loss rate between sender and receiver;

updating the rate distortion function based upon results of the encoded video object plane and upon a memory containing results of one or more previously encoded video object planes;

after the encoding the MVOs in the video object plane, performing a frame skipping function; and

transmitting, at the estimated available bandwidth, the encoded video object plane from the sender.

28. The method as defined in Claim 27, wherein:

$R_T$ is the total bit budget for the current time instant obtained from the rvcrate; and

allocating a portion of the mixed media data stream to a global buffer comprises:

$W_{cur} = \max(((W_{prev} + B_{prev}) \times R_T / R_{old} - R_T / F),\ 0)$, as the global buffer size $R_{old}/2$ is changed to $R_T/2$, wherein:

$W_{cur}$ is the occupancy of the global buffer;

$B_{prev}$ is the number of bits spent in the global buffer in the previous time instant,

$R_{old}/2$ is the previous size of the global buffer;

$W_{prev}$ is the previous occupancy of the global buffer; and

F is the video frame rate.

29. The method as defined in Claim 28, wherein allocating a portion of the mixed media data stream to a global buffer further comprises the allocation of a output target rate from the global buffer among each of video and audio data streams so as to yield the target bits for an individual object in the data stream.

30. The method as defined in Claim 28, wherein updating the rate distortion function based upon results of the encoded video object plane and upon a memory containing results of one or more previously encoded video object planes comprises a minimization of $D = \sum_i \alpha_i \times d_i$, subject to $R = \sum_i r_i \leq R_T$; wherein:

$R_T = \sum_i currate_i$,

the sending rate of the ith media stream is $r_i$,

the distortion of ith media as $d_i$, and the quality-impact parameter of the data stream is $\alpha_i$.

31. The method as defined in Claim 27, further comprising:

receiving the encoded video object plane at the receiver from the connection;

demultiplexing the encoded video object plane into coded video and auto streams;

inputting the coded video and audio streams, respectively, into video and audio decoders;

inputting the decoded video and audio streams to a media mixer; and

inputting the mixed video and audio streams output from the media mixer to an output device.

32. A computer-readable media comprising computer-executable instructions for performing the method as recited in claim 19.

33. One or more computer-readable media, comprising stored thereon:

a first set of elements to describe a server in communication through a connection of a packet switched network to a client;

a second set of elements to describe the transmission of a mixed media data stream, including audio and multiple video objects (MVOs), from the server to the client through a connection over a packet switched network; and

a code segment that, when executed:

monitors transmission characteristics of the connection between server and receiver;

estimates available bandwidth at the sender based upon the monitored transmission characteristics of the connection;

allocates a global buffer for the mixed media data stream to be transmitted stream from the sender to the receiver as a function of the estimated available bandwidth at the sender;

pre-encodes a portion of each Video Object Plane (VOP) in the global buffer with respect to a quantization parameter (QP) of the VOP;

encodes the VOP in the global buffer based on the QP;

updates a rate distortion model based upon the QP and packet loss rate;

performs a frame skipping function after the VOP encoding; and

transmits from the sender to the receiver the encoded video object plane in the global buffer at a regulated sender transmission rate from the sender as a function of the estimated available bandwidth at the sender.

34. One or more computer-readable media, comprising stored thereon:

a first set of elements to describe a server in communication through a connection of a packet switched network to a client;

a second set of elements to describe the transmission of a mixed media data stream, including audio and multiple video objects (MVOs), from the server to the client through a connection over a packet switched network; and

a code segment that, when executed:

monitors transmission characteristics of one or more encoded video object planes through the connection between the sender and the receiver;

estimates, from the transmission characteristics, an available bandwidth ($R_T$) at the sender;

allocates, as a function of the $R_T$, a portion of the mixed media data stream to a global buffer;

encodes a video object pane from the global buffer based upon a rate distortion function that accounts for packet loss rate between sender and receiver;

updates the rate distortion function based upon results of the encoded video object plane and upon a memory containing results of one or more previously encoded video object panes;

after the encoding the MVOs in the video object plane, performs a frame skipping function; and

transmits, at the estimated available bandwidth, the encoded video object plane from the sender to the receiver.